

On the Fourier transforms of self-similar measures

Masato TSUJII
Department of Mathematics,
Kyushu University

December 10, 2012

1 Introduction

For contractive affine maps $S_i : \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, 2, \dots, m$ and probabilities $0 < p_i < 1$ with $\sum_{i=1}^m p_i = 1$, there exists a unique probability measure μ such that

$$\mu = \sum_{i=1}^m p_i \cdot \mu \circ S_i^{-1}.$$

The probability measure μ thus defined is called *self-similar measure* and has been studied extensively in fractal geometry[1]. In this paper, we pose a few questions about the Fourier transforms of self-similar measures and then give a partial answer to one of them. To explain the questions, let us consider a simple setting:

$$m = 2, \quad S_1(x) = x/3, \quad S_2(x) = (x+2)/3, \quad p_1 = p_2 = 1/2,$$

in which the self-similar measure μ is the (normalized) Hausdorff measure of dimension

$$\delta = \log 2 / \log 3$$

on the standard middle-third Cantor set in $[0, 1]$. It is not difficult to see that the Fourier transform of the measure μ is given as the infinite product

$$\mathcal{F}\mu(\xi) = \prod_{\ell=1}^{\infty} \left(\frac{1 + \exp(2 \cdot 3^{-\ell} \sqrt{-1} \cdot \xi)}{2} \right). \quad (1)$$

A well-known result of Strichartz[2], which holds in more general setting, gives the asymptotic formula

$$\frac{1}{2R} \int_{-R}^R |\mathcal{F}\mu(\xi)|^2 d\xi = \mathcal{O}(R^{-\delta}) \quad \text{as } R \rightarrow \infty.$$

This implies roughly that the square-value $|\mathcal{F}\mu(\xi)|^2$ decreases like $|\xi|^{-\delta}$ as $|\xi|$ tends to infinity if it is averaged appropriately. We are interested in the deviation of the values of $\mathcal{F}\mu(\xi)$ from that average.

Let us take logarithm of the both sides of (1). Then we see

$$\log |\mathcal{F}\mu(\xi)| = \sum_{\ell=1}^{\infty} \psi(3^{-\ell}\xi)$$

where

$$\psi(x) = \log \frac{|1 + e^{2\sqrt{-1}x/3}|}{2}.$$

Note that the sum on the right side above converges because, for each ξ , the term $\psi(3^{-\ell}\xi)$ converges to 0 exponentially fast as $\ell \rightarrow \infty$. By a slightly more precise consideration, we see that there exists a constant $C > 0$, independent of N , such that

$$\left| \log |\mathcal{F}\mu(\xi)| - \sum_{\ell=1}^N \psi(3^{-\ell}\xi) \right| = \left| \sum_{\ell=N+1}^{\infty} \psi(3^{-\ell}\xi) \right| < C$$

holds for all ξ with $|\xi| \leq 3^N\pi$. Therefore we may regard the function

$$\sum_{\ell=1}^N \psi(3^{-\ell}\xi) \tag{2}$$

as an approximation of the function $\log |\mathcal{F}\mu(\xi)|$ on $[-3^N\pi, 3^N\pi]$. By change of variable $\theta = 3^{-N}\xi$, we see that the (normalized) distribution of the values of the function (2) on $[-3^N\pi, 3^N\pi]$ is same as that of the function

$$X_N : [-\pi, \pi] \rightarrow \mathbb{R}, \quad X_N(\theta) = \sum_{\ell=0}^{N-1} \psi(3^\ell\theta).$$

The average of $X_N(\cdot)/N$ is

$$A := \frac{1}{N \cdot 2\pi} \int_{-\pi}^{\pi} X_N(\theta) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \psi(\xi) d\xi < \log \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|1 + e^{2\sqrt{-1}\xi/3}|}{2} d\xi \right).$$

To analyze the deviation of the values of $X_N(\cdot)/N$ from the average A , it is useful to introduce a view point of dynamical systems. Let us consider the dynamical system generated by the map

$$T : \mathbb{R}/2\pi\mathbb{Z} \rightarrow \mathbb{R}/2\pi\mathbb{Z}, \quad T(x) = 3x \mod 2\pi\mathbb{Z}$$

and view the function $X_N(\theta)/N$ as the average of the observable ψ along the orbit of θ :

$$\frac{1}{N} X_N(\theta) = \frac{1}{N} \sum_{\ell=0}^{N-1} \psi \circ T^\ell(\theta).$$

Then the results of Takahashi[4] and Young[3] on the large deviation principle in dynamical system setting tells that¹, for any $c \geq 0$, it holds

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \left(\frac{m\{x \in [-3^N \pi, 3^N \pi] \mid |\mathcal{F}\mu(\xi)| \geq e^{-cN}\}}{3^N \cdot 2\pi} \right) \leq \widehat{R}(c) - \log 3$$

with setting

$$\widehat{R}(c) = \sup \left\{ h_\nu(T) \mid \nu \in \mathcal{M}_T \text{ such that } \int \psi d\nu \geq -c \right\} \geq 0$$

where \mathcal{M}_T denotes the space of T -invariant probability measures and $h_\nu(T)$ the measure theoretical entropy of T with respect to a measure $\nu \in \mathcal{M}_T$.

The function $\widehat{R}(c)$ is increasing by definition. Further we can make the following observations: (Figure 1)

- (A) $\widehat{R}(c) - \log 3 \leq 0$ and the inequality is strict if and only if $c < |A| = -\int_{-\pi}^{\pi} \psi(\xi) d\xi$, because the (normalized) uniform measure is the unique maximal entropy ($= \log 3$) measure for T .
- (B) $\lim_{c \rightarrow 0} \widehat{R}(c) = 0$. This is because, if $c > 0$ is close to 0, a T -invariant probability measure ν satisfying $\int \psi d\nu \geq -c$ must concentrate mostly on a small neighborhood of the fixed point $0 \in \mathbb{R}/2\pi\mathbb{Z}$.
- (C) The function $\widehat{R}(c)$ is a convex function on $[0, |A|]$. This follows immediately from the definition of $\widehat{R}(c)$ and the affine property of the entropy $h_\nu(T)$ with respect to ν .

The question that we would like to pose is whether and/or how these facts remains true for more general classes of self-similar measures (or classes of dynamically defined measures). It seems that the observations above holds true with obvious modifications if the contraction rates of the similarities are same. But, if the contraction rates are different, it is not clear that whether the observations still hold or not. In the following, we give a simple (and rather modest) result related to the observation (B).

2 Result

For $i = 1, 2, \dots, m$, let

$$S_i : \mathbb{R} \rightarrow \mathbb{R}, \quad S_i(x) = a_i x + b_i$$

¹Actually we can not apply the results in [4] and [3] directly because the function ψ has logarithmic singularities and hence is not continuous. Still we can obtain the large deviation estimate stated here by approximating ψ by smooth functions from above. It seems reasonable to expect that the limit is exact and the inequality is actually an equality.

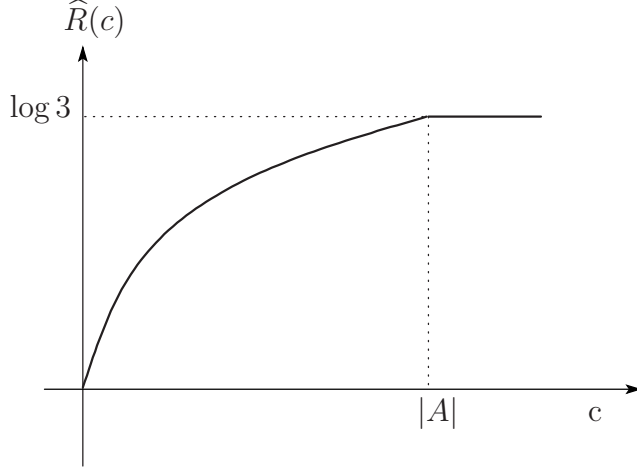


Figure 1: A schematic picture of the graph of $\hat{R}(c)$

be contracting affine maps. Let $0 < p_i < 1$, $i = 1, 2, \dots, m$, be real numbers such that $\sum_{i=1}^m p_i = 1$. Then there exists a unique probability measure μ such that

$$\mu = \sum_{i=1}^m p_i \cdot \mu \circ S_i^{-1}.$$

The support K of the measure μ is the unique compact subset that satisfies

$$K = \bigcup_{i=1}^m S_i(K).$$

We henceforth assume that the subset K is not a single point, to avoid the trivial case. The Fourier transform $\mathcal{F}\mu$ of the measure μ is a real-analytic function defined by

$$\mathcal{F}\mu(\xi) = \int \exp(-\sqrt{-1}\xi x) d\mu(x).$$

Let $R(c; \mu)$ be the function defined for μ by

$$R(c; \mu) = \limsup_{t \rightarrow \infty} \frac{1}{t} \log (\text{Leb}\{x \in [-e^t, e^t] \mid |\mathcal{F}\mu(\xi)| \geq e^{-ct}\}). \quad (3)$$

The following is our result, which tells that the observation (B) in the last section holds generally for (non-trivial) self-similar measures.

Theorem. *For any $\epsilon > 0$, there exists a positive constant $C(\epsilon) > 0$ such that*

$$R(c; \mu) \leq \epsilon + C(\epsilon) \cdot c \cdot |\log c|$$

for $c > 0$. In particular, we have $\lim_{c \rightarrow +0} R(c; \mu) = 0$.

3 Proof

3.1 Preliminaries

Let $\mathcal{A} = \{1, 2, \dots, m\}$. Let \mathbf{p} be the probability measure on \mathcal{A} and set $p_i = \mathbf{p}(\{i\})$. We denote by \mathcal{A}^n the set of words of length n with letters in \mathcal{A} and by \mathbf{p}^n the probability measure on it obtained as the n -times direct product of \mathbf{p} .

We regard the system $\{(S_i, p_i)\}_{i \in \mathcal{A}}$ as a random dynamical system in which the mapping S_i is applied with probability p_i . Then the measure μ in the main theorem is the unique stationary measure for it. The n -th iteration of the random dynamical system $\{(S_i, p_i)\}_{i \in \mathcal{A}}$ is to the system $\{(S_{\mathbf{i}}, p_{\mathbf{i}})\}_{\mathbf{i} \in \mathcal{A}^n}$ where

$$S_{\mathbf{i}} : \mathbb{R} \rightarrow \mathbb{R}, \quad S_{\mathbf{i}} = S_{i(n)} \circ S_{i(n-1)} \circ \dots \circ S_{i(1)}$$

and

$$p_{\mathbf{i}} = p_{i(1)} \cdot p_{i(2)} \cdot \dots \cdot p_{i(n)}$$

for $\mathbf{i} = (i(1), i(2), \dots, i(n))$. The affine map $S_{\mathbf{i}}$ is expressed as

$$S_{\mathbf{i}}(x) = a_{\mathbf{i}}x + b_{\mathbf{i}}$$

where

$$a_{\mathbf{i}} = \prod_{k=1}^n a_{i(k)}.$$

Note that the measure μ is the unique invariant measure also for the iterations as above. Below we develop our argument for the system $\{(S_i, p_i)\}_{i \in \mathcal{A}}$. But our argument is applicable to the iterations of $\{(S_i, p_i)\}_{i \in \mathcal{A}}$ in parallel and, in a few places, we will replace the system $\{(S_i, p_i)\}_{i \in \mathcal{A}}$ by its iterate in order to assume some numerical conditions. For instance, we may (and will) assume

$$|a_i| < \frac{1}{2} \quad \text{for all } 1 \leq i \leq m \tag{4}$$

without loss of generality, by such replacement.

Any changes of coordinate on \mathbb{R} by affine bijections do not affect validity of the main theorem. Therefore we may and do assume also

$$b_0 = 0 \leq b_1 \leq \dots \leq b_m = 1. \tag{5}$$

It is then easy to see that the invariant subset K is contained in $[-2, 2]$.

3.2 The average rate of contraction

The averaged Lyapunov exponent of the random dynamical system $\{S_i, p_i\}_{i=1}^m$ is

$$\chi := \sum_{i=1}^m p_i \cdot \log |a_i| < 0.$$

The following is a simple application of the large deviation principle[5].

Lemma 1. *For any $\delta > 0$, there exists $\epsilon > 0$ such that*

$$\mathbf{p}^n \{\mathbf{i} \in \mathcal{A}^n \mid |a_{\mathbf{i}}| \leq \exp((\chi - \delta)n)\} \leq e^{-\epsilon n}$$

and

$$\mathbf{p}^n \{\mathbf{i} \in \mathcal{A}^n \mid |a_{\mathbf{i}}| \geq \exp((\chi + \delta)n)\} \leq e^{-\epsilon n}$$

hold for sufficiently large $n \geq 0$.

Take and fix integers $0 < \check{r} < \hat{r}$ with $\hat{r} - \check{r} \leq 2$ so that

$$\check{r} < e^{|\chi|} < \hat{r}.$$

Corollary 2. *There exists a constant $\epsilon > 0$ such that*

$$\mathbf{p}^n \{\mathbf{i} \in \mathcal{A}^n \mid |a_{\mathbf{i}}| \geq \check{r}^{-n} \text{ or } |a_{\mathbf{i}}| \leq \hat{r}^{-n}\} \leq e^{-\epsilon n}$$

for sufficiently large $n \geq 0$.

Notice that, replacing $\{S_i, p_i\}_{i=1}^m$ by its iteration, we may assume that the integers \hat{r} and \check{r} are arbitrarily large and that the rate \hat{r}/\check{r} is arbitrarily close to 1.

3.3 The sequence of measures μ_n and their Fourier transform

Let δ_0 be the Dirac measure at the origin $0 \in \mathbb{R}$ and set

$$\mu_n = \sum_{\mathbf{i} \in \mathcal{A}^n} p_{\mathbf{i}} \cdot \delta_0 \circ S_{\mathbf{i}}^{-1} \quad \text{for } n \geq 0.$$

The measure μ_n converges to μ as $n \rightarrow \infty$ in the weak topology and hence its Fourier transform $\mathcal{F}\mu_n(\xi)$ converges to $\mathcal{F}\mu(\xi)$ for each $\xi \in \mathbb{R}$. The convergence is actually uniform on any compact subset. More precisely we have

Lemma 3. *There exists a constant $\epsilon > 0$ such that, for sufficiently large n , we have*

$$|\mathcal{F}\mu(\xi) - \mathcal{F}\mu_n(\xi)| \leq e^{-\epsilon n} \quad \text{uniformly for } \xi \in \mathbb{R}^d \text{ with } |\xi| \leq \check{r}^n.$$

Proof. From invariance of μ , we have

$$\mathcal{F}\mu(\xi) - \mathcal{F}\mu_n(\xi) = \sum_{\mathbf{i} \in \mathcal{A}^n} p_{\mathbf{i}} \cdot \int e^{-\sqrt{-1}\xi \cdot x} d((\delta_0 - \mu) \circ S_{\mathbf{i}}^{-1})(x). \quad (6)$$

Take a real number r in between \check{r} and $e^{|\chi|}$. From Lemma 1, we see that

$$\mathbf{p}^n \{\mathbf{i} \in \mathcal{A}^n \mid |a_{\mathbf{i}}| \geq r^{-n}\}$$

is exponentially small. That is, the sum in (6) restricted to $\mathbf{i} \in \mathcal{A}^n$ with $|a_{\mathbf{i}}| \geq r^{-n}$ is exponentially small with respect to n . If $|a_{\mathbf{i}}| < r^{-n}$, we have that

$$\left| \int e^{-\sqrt{-1}\xi \cdot x} d((\delta_0 - \mu) \circ S_{\mathbf{i}}^{-1})(x) \right| < |\xi| \cdot r^{-n} \cdot \text{diam} K < 4(\check{r}/r)^n$$

provided $|\xi| \leq \check{r}^n$. Therefore the remaining part of the sum (6) is also exponentially small with respect to n . \square

To continue, we introduce the operators

$$T_i : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R}), \quad T_i u(\xi) = e^{-\sqrt{-1}b_i \xi} \cdot u(a_i \cdot \xi)$$

defined for $i \in \mathcal{A}$. Since $\mathcal{F}(\nu \circ S_i^{-1}) = T_i(\mathcal{F}\nu)$, we have

$$\mathcal{F}\mu_n = \sum_{\mathbf{i} \in \mathcal{A}^n} p_{\mathbf{i}} \cdot T_{\mathbf{i}}(\mathbf{1})$$

where $T_{\mathbf{i}} = T_{\mathbf{i}(n)} \circ T_{\mathbf{i}(n-1)} \circ \cdots \circ T_{\mathbf{i}(1)}$.

Let n be a large positive integer. For $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n$, let $X_{\mathbf{i}} : I(\mathbf{i}) \rightarrow \mathbb{C}$ be the restriction of the function $\mathcal{F}\mu_{n-k}(\xi)$ to the interval

$$I(\mathbf{i}) := [-|a_{\mathbf{i}}| \cdot \check{r}^n, |a_{\mathbf{i}}| \cdot \check{r}^n].$$

Then, from the definition, we have that

$$X_{\mathbf{i}}(\xi) \equiv 1 \quad \text{for all } \mathbf{i} \in \mathcal{A}^n \quad (7)$$

and also that

$$X_{\emptyset}(\xi) = \mathcal{F}\mu_n(\xi).$$

The functions $X_{\mathbf{i}}(\xi)$ for $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n-1$ are in the relation

$$X_{\mathbf{i}}(\xi) = \sum_{j=1}^m p_j \cdot T_j X_{j \cdot \mathbf{i}}(\xi) = \sum_{j=1}^m p_j \cdot \exp(-\sqrt{-1}b_j \xi) \cdot X_{j \cdot \mathbf{i}}(a_j \xi) \quad (8)$$

where $j \cdot \mathbf{i} \in \mathcal{A}^{k+1}$ is the sequence defined by

$$j \cdot \mathbf{i}(\ell) = \begin{cases} j, & \text{if } \ell = 1; \\ \mathbf{i}(\ell - 1), & \text{if } 2 \leq \ell \leq k + 1. \end{cases}$$

3.4 The auxiliary quantity $Y_{\mathbf{i}}(\cdot)$

Our task is to analyze how the summands on the right hand side of (8) cancel each other by the difference of the complex phase. One technical problem in such analysis is that, if a few of the summands are much larger than others in absolute value, the cancellation will not happen effectively. In order to deal with such problem, we introduce another family of positive real-valued functions

$$Y_{\mathbf{i}} : I(\mathbf{i}) \rightarrow \mathbb{R}_+$$

for $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n$. For $\mathbf{i} \in \mathcal{A}^n$, we set

$$Y_{\mathbf{i}}(\xi) = X_{\mathbf{i}}(\xi) \equiv 1 \quad \text{on } I(\mathbf{i}).$$

Then we define $Y_{\mathbf{i}}(\xi)$ inductively by the relation

$$Y_{\mathbf{i}}(\xi) = \max \left\{ \frac{1}{2} \sum_{j=1}^m p_j \cdot Y_{j,\mathbf{i}}(a_j \cdot \xi), |X_{\mathbf{i}}(\xi)| \right\}.$$

From this definition and the relation (8), we have that

$$|X_{\mathbf{i}}(\xi)| \leq Y_{\mathbf{i}}(\xi) \leq 1$$

and that

$$\frac{1}{2} \sum_{j=1}^m p_j \cdot Y_{j,\mathbf{i}}(a_j \cdot \xi) \leq Y_{\mathbf{i}}(\xi) \leq \sum_{j=1}^m p_j \cdot Y_{j,\mathbf{i}}(a_j \cdot \xi). \quad (9)$$

The next lemma is the main reason to introduce the functions $Y_{\mathbf{i}}(\cdot)$.

Lemma 4. *For any $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n$, we have*

$$\left| \frac{d}{d\xi} X_{\mathbf{i}}(\xi) \right| \leq \frac{1}{1 - 2(\max_{1 \leq i \leq m} |a_i|)} \cdot Y_{\mathbf{i}}(\xi) \quad \text{for any } \xi \in I(\mathbf{i})$$

and

$$\left| \frac{d}{d\xi} Y_{\mathbf{i}}(\xi) \right| \leq \frac{1}{1 - 2(\max_{1 \leq i \leq m} |a_i|)} \cdot Y_{\mathbf{i}}(\xi) \quad \text{for Lebesgue almost all } \xi \in I(\mathbf{i})$$

where the differential in the second inequality is considered in the sense of distribution. (Note that $Y_{\mathbf{i}}$ satisfies the Lipschitz condition and hence $\frac{d}{d\xi} Y_{\mathbf{i}} \in L^\infty$.)

Proof. By the chain rule and the assumption $|b_i| \leq 1$, we have

$$\left| \frac{d}{d\xi} X_{\mathbf{i}}(\xi) \right| \leq \sum_{\ell=0}^{n-k} \sum_{\mathbf{j} \in \mathcal{A}^\ell} p_{\mathbf{j}} \cdot |a_{\mathbf{j}}| \cdot |X_{\mathbf{j},\mathbf{i}}(a_{\mathbf{j}} \cdot \xi)| \leq \sum_{\ell=0}^{n-k} \sum_{\mathbf{j} \in \mathcal{A}^\ell} p_{\mathbf{j}} \cdot |a_{\mathbf{j}}| \cdot Y_{\mathbf{j},\mathbf{i}}(a_{\mathbf{j}} \cdot \xi).$$

From the assumption (4) and the left inequality in (9), the sum on the right hand side is bounded by

$$\begin{aligned} \sum_{\ell=0}^{n-k} \left(\sum_{\mathbf{j} \in \mathcal{A}^\ell} 2^\ell |a_{\mathbf{j}}| \cdot 2^{-\ell} \cdot p_{\mathbf{j}} \cdot Y_{\mathbf{j},\mathbf{i}}(a_{\mathbf{j}} \cdot \xi) \right) &\leq \sum_{\ell=0}^{n-k} (2 \max_{1 \leq i \leq m} |a_i|)^\ell \cdot \left(2^{-\ell} \cdot \sum_{\mathbf{j} \in \mathcal{A}^\ell} p_{\mathbf{j}} \cdot Y_{\mathbf{j},\mathbf{i}}(a_{\mathbf{j}} \cdot \xi) \right) \\ &\leq \frac{1}{1 - 2(\max_{1 \leq i \leq m} |a_i|)} \cdot Y_{\mathbf{i}}(\xi). \end{aligned}$$

We thus obtained the first claim of the lemma. We can get the second claim by inductive argument on the length of \mathbf{i} , using the first claim. \square

3.5 The main estimate

For a sequence $\mathbf{i} = (\mathbf{i}(1), \mathbf{i}(2), \dots, \mathbf{i}(k)) \in \mathcal{A}^k$ and $1 \leq \ell \leq \ell' \leq k$, we define

$$[\mathbf{i}]_{\ell}^{\ell'} = (\mathbf{i}(\ell), \mathbf{i}(\ell+1), \dots, \mathbf{i}(\ell')) \in \mathcal{A}^{\ell'-\ell+1}.$$

For $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n$, let $\Xi_{\mathbf{i}}$ be the partition of the interval $[-\tilde{r}^n, \tilde{r}^n]$ defined as follows: For $\mathbf{i} = \emptyset \in \mathcal{A}^0$, let Ξ_{\emptyset} be the partition of the interval $I(\emptyset) = [-\tilde{r}^n, \tilde{r}^n]$ into the unit intervals

$$[k, k+1] \quad \text{for } -\tilde{r}^n \leq k < \tilde{r}^n.$$

Then we construct the partition $\Xi_{\mathbf{i}}$ for $\mathbf{i} \in \mathcal{A}^k$ inductively so that each element of $\Xi_{\mathbf{i}}$ is a union of consecutive intervals in the partition $\Xi_{\mathbf{i}'}$ with $\mathbf{i}' = [\mathbf{i}]_2^k \in \mathcal{A}^{k-1}$ such that

$$\min\{|a_{\mathbf{i}}|^{-1}, 2\tilde{r}^n\} \leq |I| \leq 2|a_{\mathbf{i}}|^{-1} \quad \text{for } I \in \Xi_{\mathbf{i}}.$$

Such partition is not unique of course. But any of such partition will work in the following argument. Note that, when $|a_{\mathbf{i}}|^{-1} \geq 2\tilde{r}^n$, the partition $\Xi_{\mathbf{i}}$ consists of the single interval $[-\tilde{r}^n, \tilde{r}^n]$ and, otherwise, we have

$$|a_{\mathbf{i}}|^{-1} \leq |I| \leq 2|a_{\mathbf{i}}|^{-1} \quad \text{for } I \in \Xi_{\mathbf{i}}.$$

The next lemma deals with the technical part of the proof of Theorem 2.

Lemma 5. *There exists $\eta > 0$, independent of n , such that the following holds: Consider arbitrary $j \in \mathcal{A}$, $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k \leq n-1$ and $I \in \Xi_{j, \mathbf{i}}$. From the construction of the partition $\Xi_{j, \mathbf{i}}$ explained above, the interval I is the union of consecutive intervals $I_1, I_2, \dots, I_N \in \Xi_{\mathbf{i}}$. Set*

$$I'_i := a_{j, \mathbf{i}} \cdot I_i \subset I' := a_{j, \mathbf{i}} \cdot I \quad \text{for } 1 \leq i \leq N. \quad (10)$$

Then we have

$$Y_{\mathbf{i}}(\xi)^2 \leq \exp(-\eta) \cdot \sum_{\ell \in \mathcal{A}} p_{\ell} \cdot Y_{\ell, \mathbf{i}}(a_{\ell} \cdot \xi)^2 \quad \text{for } \xi \in I'_i \quad (11)$$

for all $1 \leq i \leq N$ but for at most two exceptions.

Proof. We may and do assume $N > 2$ because the claim is vacuous otherwise. This implies $|a_{\mathbf{i}}|^{-1} < 2\tilde{r}^n$ as we noted above. In particular, we have

$$|I| \leq 2|a_{j, \mathbf{i}}|^{-1} \quad \text{and} \quad |a_{\mathbf{i}}|^{-1} \leq |I_i| \leq 2|a_{\mathbf{i}}|^{-1}$$

and hence

$$|I'| \leq 2 \quad \text{and} \quad |a_j| \leq |I'_i| \leq 2|a_j| \leq 1. \quad (12)$$

We first show that the claim of the lemma can be proved by simple argument if either of the following two conditions holds:

(I) $Y_{\ell, \mathbf{i}}(a_{\ell} \cdot \xi) \leq Y_{\ell', \mathbf{i}}(a_{\ell'} \cdot \xi)/2$ on I' for some pair $(\ell, \ell') \in \mathcal{A}$.

(II) $X_{\ell_0, \mathbf{i}}(a_{\ell_0} \cdot \xi) \leq Y_{\ell_0, \mathbf{i}}(a_{\ell_0} \cdot \xi)/2$ on I' for some $\ell_0 \in \mathcal{A}$.

Suppose that the condition (I) holds. Setting $a_\ell = \sqrt{p_\ell}$ and $b_\ell = \sqrt{p_\ell} \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)$ in the identity

$$\left(\sum_{\ell \in \mathcal{A}} a_\ell b_\ell \right)^2 = \left(\sum_{\ell \in \mathcal{A}} a_\ell^2 \right) \left(\sum_{\ell \in \mathcal{A}} b_\ell^2 \right) - \frac{1}{2} \sum_{\ell, \ell' \in \mathcal{A}} (a_\ell b_{\ell'} - a_{\ell'} b_\ell)^2,$$

we get

$$\left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi) \right)^2 = \sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)^2 - \sum_{\ell, \ell' \in \mathcal{A}} p_\ell p_{\ell'} \cdot \frac{|Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi) - Y_{\ell', \mathbf{i}}(a_{\ell'} \cdot \xi)|^2}{2}.$$

Hence we have, from (9), that

$$1 - \frac{Y_{\mathbf{i}}(\xi)^2}{\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)^2} \geq \frac{p_{\min}^2 \cdot |Y_{\max}(\xi) - Y_{\min}(\xi)|^2}{2 \cdot \sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)^2} \geq \frac{p_{\min}^2 \cdot |Y_{\max}(\xi) - Y_{\min}(\xi)|^2}{2 \cdot Y_{\max}(\xi)^2}$$

with setting $p_{\min} = \min_{i \in \mathcal{A}} p_i > 0$ and

$$Y_{\max}(\xi) = \max_{\ell \in \mathcal{A}} Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi), \quad Y_{\min}(\xi) = \min_{\ell \in \mathcal{A}} Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi).$$

From the condition (I), the right hand side of the inequality above is not less than $p_{\min}^2/8$. Therefore, setting $\eta = p_{\min}^2/8$, we obtain the conclusion of the lemma (with no exception for $1 \leq i \leq N$).

Next suppose that the condition (II) holds and that the condition (I) does *not* hold. From the latter condition, there exists a point $\xi_0 \in I'$ for each pair $(\ell, \ell') \in \mathcal{A}^2$ such that

$$Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi_0) > Y_{\ell', \mathbf{i}}(a_{\ell'} \cdot \xi_0)/2. \quad (13)$$

From Lemma 4 and (10), we have

$$\frac{Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)}{Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi')} \leq \alpha := \exp \left(\frac{2 \cdot \max_{1 \leq i \leq m} |a_i|}{1 - 2(\max_{1 \leq i \leq m} |a_i|)} \right) \quad (14)$$

for all $\xi, \xi' \in I'$ and $\ell \in \mathcal{A}$. Hence

$$Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi) \geq \frac{1}{2\alpha^2} \cdot Y_{\ell', \mathbf{i}}(a_{\ell'} \cdot \xi) \quad \text{for all } \xi \in I' \text{ and } (\ell, \ell') \in \mathcal{A}^2. \quad (15)$$

In particular, setting $c = p_{\min}/(2\alpha^2)^2$, we have

$$p_\ell \cdot Y_{\ell, \mathbf{i}}(a_\ell \cdot \xi)^2 \geq c \cdot \sum_{\ell' \in \mathcal{A}} p_{\ell'} \cdot Y_{\ell', \mathbf{i}}(a_{\ell'} \cdot \xi)^2 \quad \text{for all } \xi \in I' \text{ and } \ell \in \mathcal{A}. \quad (16)$$

By the definition of $Y_i(\xi)$ and Schwarz inequality, we have

$$\begin{aligned}
Y_i(\xi)^2 &\leq \max \left\{ \frac{1}{4} \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell,i}(a_\ell \cdot \xi) \right)^2, \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot |X_{\ell,i}(a_\ell \cdot \xi)| \right)^2 \right\} \\
&\leq \max \left\{ \frac{1}{4} \sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell,i}(a_\ell \cdot \xi)^2, \sum_{\ell \in \mathcal{A}} p_\ell \cdot |X_{\ell,i}(a_\ell \cdot \xi)|^2 \right\} \\
&\leq \sum_{\ell \in \mathcal{A}} p_\ell \cdot \max \left\{ \frac{1}{4} Y_{\ell,i}(a_\ell \cdot \xi)^2, |X_{\ell,i}(a_\ell \cdot \xi)|^2 \right\}
\end{aligned}$$

for all $\xi \in I'$. But, from the condition (II), the last expression is bounded by

$$\left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell,i}(a_\ell \cdot \xi)^2 \right) - \frac{1}{2} p_{\ell_0} \cdot Y_{\ell_0,i}(a_{\ell_0} \cdot \xi)^2$$

where $\ell_0 \in \mathcal{A}$ is that in the condition (II). This together with the estimate (16) implies the conclusion of the lemma (again with no exception for $1 \leq i \leq N$), setting $\eta = c/2$.

Below we consider the remaining case where neither of the conditions (I) and (II) holds. Notice that the estimates (14), (15) and (16) still hold in this case, because these are consequences of the assumption that the condition (I) does not hold. From the assumption that the condition (II) does not hold, for each $\ell \in \mathcal{A}$, there exists a point $\xi_0 \in I$ such that

$$|X_{\ell,i}(\xi_0)| \geq Y_{\ell,i}(\xi_0)/2.$$

From Lemma 4 and (14), we have that

$$\left| \frac{d}{d\xi} (X_{\ell,i}(a_\ell \cdot \xi)) \right| \leq \frac{\max_{1 \leq i \leq m} |a_i|}{1 - 2 \cdot \max_{1 \leq i \leq m} |a_i|} \cdot Y_{\ell,i}(\xi) \leq \frac{\max_{1 \leq i \leq m} |a_i|}{1 - 2 \cdot \max_{1 \leq i \leq m} |a_i|} \cdot \alpha \cdot Y_{\ell,i}(\xi_0)$$

for all $\xi \in I'$ and $\ell \in \mathcal{A}$. As we noted in the beginning, by replacing the system $\{(S_i, p_i)\}_{i \in \mathcal{A}}$ by its iterates, we may and do assume that the constant α in (14) is close to 1 and that

$$\frac{\max_{1 \leq i \leq m} |a_i|}{1 - 2 \cdot \max_{1 \leq i \leq m} |a_i|}$$

is close to 0. Thus we may suppose without loss of generality that the inequality above implies that

$$|X_{\ell,i}(a_\ell \cdot \xi)| \geq Y_{\ell,i}(a_\ell \cdot \xi')/8 \tag{17}$$

and

$$\left| \frac{d}{d\xi} (X_{\ell,i}(a_\ell \cdot \xi)) \right| \leq \frac{|X_{\ell,i}(a_\ell \cdot \xi')|}{10} \tag{18}$$

hold for all $\xi, \xi' \in I'$ and $\ell \in \mathcal{A}$.

Now we make use of the difference of the complex argument of the summands on the right hand side of (8). From the definition of $Y_i(\cdot)$ and the relation (8), we have that

$$\begin{aligned} \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, i}(a_\ell \cdot \xi) \right)^2 - |X_i(\xi)|^2 &\geq \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot |T_\ell X_{\ell, i}(\xi)| \right)^2 - \left| \sum_{\ell \in \mathcal{A}} p_\ell \cdot T_\ell X_{\ell, i}(\xi) \right|^2 \\ &\geq (1 - \cos \Theta_{\ell' \ell''}(\xi)) \cdot p_{\ell'} \cdot p_{\ell''} \cdot |T_{\ell'} X_{\ell', i}(\xi)| \cdot |T_{\ell''} X_{\ell'', i}(\xi)| \end{aligned}$$

where ℓ' and ℓ'' are arbitrary two distinct element in \mathcal{A} and $\Theta_{\ell' \ell''}(\xi)$ denotes the difference between the complex arguments of $T_{\ell'} X_{\ell', i}(\xi)$ and $T_{\ell''} X_{\ell'', i}(\xi)$. From (15) and (17), we have

$$p_{\ell'} \cdot p_{\ell''} \cdot |T_{\ell'} X_{\ell', i}(\xi)| \cdot |T_{\ell''} X_{\ell'', i}(\xi)| \geq c' \cdot \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, i}(a_\ell \cdot \xi) \right)^2,$$

setting $c' = (1/8)^2 \cdot (p_{\min}/(2\alpha^2))^2 > 0$. Therefore, for all $\xi \in I'$, we have

$$|X_i(\xi)|^2 \leq (1 - c'(1 - \cos \Theta_{\ell' \ell''}(\xi))) \cdot \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, i}(a_\ell \cdot \xi) \right)^2$$

and hence

$$\begin{aligned} |Y_i(\xi)|^2 &\leq \min\{1/4, 1 - c'(1 - \cos \Theta_{\ell' \ell''}(\xi))\} \cdot \left(\sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, i}(a_\ell \cdot \xi) \right)^2 \\ &\leq (1 - c'(1 - \cos \Theta_{\ell' \ell''}(\xi))) \cdot \sum_{\ell \in \mathcal{A}} p_\ell \cdot Y_{\ell, i}(a_\ell \cdot \xi)^2. \end{aligned}$$

Let us set $\ell' = 1$ and $\ell'' = m$ so that $b_{\ell'} - b_{\ell''} = 1$ from the assumption (5). Then, from (18) and the definition of the operator T_ℓ , we have that

$$\left| \frac{d}{d\xi} \Theta_{\ell' \ell''}(\xi) - (b_{\ell''} - b_{\ell'}) \right| = \left| \frac{d}{d\xi} \Theta_{\ell' \ell''}(\xi) - 1 \right| \leq \frac{2}{10}.$$

Since the length of the interval I' and the subintervals I'_j satisfy (12), this implies that, for some small constant $\rho > 0$ independent of n , we have

$$\cos \Theta_{\ell' \ell''}(\xi) \leq 1 - \rho \quad \text{on } I'_j$$

for all $1 \leq j \leq N$ but for at most two exceptions. We therefore obtain the conclusion of the lemma by choosing $\eta > 0$ so small that $e^{-\eta} = 1 - c'\rho$. \square

3.6 Proof of Theorem 2

From the argument in the last section, we recall that

- from Lemma 3, we have

$$|\mathcal{F}\mu(\xi)| \leq |\mathcal{F}\mu_n(\xi)| + e^{-\epsilon n} = |X_\emptyset(\xi)| + e^{-\epsilon n} \leq Y_\emptyset(\xi) + e^{-\epsilon n}$$

for $\xi \in [-\tilde{r}^n, \tilde{r}^n]$, and that

- the functions $Y_{\mathbf{i}}$ for $\mathbf{i} \in \bigcup_{0 \leq k \leq n} \mathcal{A}^k$ satisfy the inductive estimate in Lemma 5.

From these observation and the construction of the partition $\Xi_{\mathbf{i}}$, it is not difficult to deduce Theorem 2. Below we give a proof.

Let $n > 0$ be a large integer and let $\eta > 0$ be the constant taken in the lemma. For a sequence $\mathbf{i} \in \mathcal{A}^k$ with $0 \leq k < n$ and a point $\xi \in [-\tilde{r}^n, \tilde{r}^n]$, we set

$$r_{\mathbf{i}}(\xi) = \begin{cases} \eta, & \text{if the inequality in (11) holds;} \\ 0, & \text{otherwise.} \end{cases}$$

Then, for $\xi \in [-\tilde{r}^n, \tilde{r}^n]$, we have

$$|\mathcal{F}\mu_n(\xi)|^2 \leq Y_\emptyset(\xi)^2 \leq \sum_{\mathbf{i} \in \mathcal{A}^n} \exp\left(-\sum_{k=2}^{n+1} r_{[\mathbf{i}]_k^n}(\xi)\right) \cdot p_{\mathbf{i}} \cdot Y_{\mathbf{i}}(a_{\mathbf{i}} \cdot \xi)^2 = \sum_{\mathbf{i} \in \mathcal{A}^n} \exp\left(-\sum_{k=2}^{n+1} r_{[\mathbf{i}]_k^n}(\xi)\right) \cdot p_{\mathbf{i}}.$$

Below we estimate the Lebesgue measure of the subset

$$Z_n := \{\xi \in [-\tilde{r}^n, \tilde{r}^n] \mid |\mathcal{F}\mu_n(\xi)|^2 \geq 3 \exp(-s\eta n)\} \quad (19)$$

for small $s > 0$. Let $\mathcal{E}_n \subset \mathcal{A}^n$ be the set of sequences $\mathbf{i} \in \mathcal{A}^n$ such that $|a_{\mathbf{i}}|^{-1} < 2\tilde{r}^n$. Then, by Lemma 1, we have

$$\mathbf{p}^n(\mathcal{E}_n) \leq \exp(-s\eta n)$$

provided that $s > 0$ is sufficiently small. Hence each $\xi \in Z_n$ satisfies

$$\sum_{\mathbf{i} \in \mathcal{A}^n \setminus \mathcal{E}_n} \exp\left(-\sum_{k=2}^{n+1} r_{[\mathbf{i}]_k^n}(\xi)\right) \cdot p_{\mathbf{i}} \geq 2 \exp(-s\eta n).$$

If we set

$$\mathbf{I}_{\mathbf{i}} = \int_{-\tilde{r}^n}^{\tilde{r}^n} \max\left\{0, \exp\left(-\sum_{k=2}^{n+1} r_{[\mathbf{i}]_k^n}(\xi)\right) - \exp(-s\eta n)\right\} d\xi$$

for $\mathbf{i} \in \mathcal{A}^n$, have

$$\exp(-s\eta n) \cdot \text{Leb} Z_n \leq \sum_{\mathbf{i} \in \mathcal{A}^n \setminus \mathcal{E}_n} p_{\mathbf{i}} \cdot \mathbf{I}_{\mathbf{i}}. \quad (20)$$

We are going to estimate the integral $\mathbf{I}_{\mathbf{i}}$ for $\mathbf{i} \in \mathcal{A}^n \setminus \mathcal{E}_n$ by an elementary combinatorial argument. First of all, note that, if $\mathbf{i} \notin \mathcal{E}_n$, we have $|a_{\mathbf{i}}|^{-1} \geq 2\tilde{r}^n$ and the partition $\Xi_{\mathbf{i}}$

consists of the single interval $[-\check{r}^n, \check{r}^n]$. From the construction of the partitions $\Xi_{\mathbf{i}}$, each interval in the partition $\Xi_{[\mathbf{i}]_k^n}$ with $0 \leq k \leq n-1$ consists of at most

$$\kappa := \left\lceil 2 \cdot \max_{i \in \mathcal{A}} |a_i|^{-1} \right\rceil.$$

consecutive intervals in $\Xi_{[\mathbf{i}]_{k+1}^n}$. And Lemma 5 tells that the condition $r_{[\mathbf{i}]_k^n}(\xi) = 0$ holds for points ξ in at most two intervals among them. Since the condition

$$\sum_{k=2}^{n+1} r_{[\mathbf{i}]_k^n}(\xi) \leq \nu\eta$$

implies that $r_{[\mathbf{i}]_k^n}(\xi) = 0$ holds for all $2 \leq k \leq n+1$ but for at most ν exceptions. Hence we obtain

$$\mathbf{I}_{\mathbf{i}} \leq \sum_{\nu \leq [sn]} \binom{n}{\nu} \cdot 2^{n-\nu} \cdot \kappa^\nu \cdot \exp(-\nu\eta).$$

By using Stirling's formula

$$\log m! = m \log m - m + \mathcal{O}(\log m)$$

and the inequality $\log(1+x) \leq x$ that holds for any $x \geq 0$, we bound the logarithms of the summands as

$$\begin{aligned} \log \left(\binom{n}{\nu} \cdot 2^{n-\nu} \cdot \kappa^\nu \cdot \exp((\nu+1)\eta) \right) &\leq (n-\nu) \log \left(\frac{n}{n-\nu} \right) + \nu \log \left(\frac{n}{\nu} \right) \\ &\quad + (n-\nu) \log 2 + \nu \log \kappa + \nu\eta + \mathcal{O}(\log n) \\ &\leq \nu(1 + |\log s| + \log \kappa + \eta) + n \log 2 + \mathcal{O}(\log n) \end{aligned}$$

for $\nu \leq [sn]$. Then we obtain

$$\begin{aligned} \mathbf{I}_{\mathbf{i}} &\leq \exp(n(s(1 + |\log s| + \eta + \log K) + \log 2) + \mathcal{O}(\log n)) \\ &\leq \exp(n(2s|\log s| + \log 2)) \end{aligned}$$

if $s > 0$ is sufficiently small and if n is sufficiently large. Putting this estimate in (20), we conclude

$$\log \text{Leb} Z_n \leq n(2s|\log s| + \log 2),$$

that is,

$$\frac{1}{n} \log (\text{Leb} \{x \in [-\check{r}^n, \check{r}^n] \mid |\mathcal{F}\mu_n(\xi)| \geq 3e^{-s\eta n}\}) \leq 2s|\log s| + \log 2$$

for sufficiently large n . Together with Lemma 3, this implies

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log (\text{Leb} \{x \in [-e^t, e^t] \mid |\mathcal{F}\mu(\xi)| \geq e^{-ct}\}) \leq \frac{2s|\log s| + \log 2}{\log \check{r}}$$

with $c = (s \cdot \eta)/(2 \log \check{r})$. Since we may assume that \check{r} is arbitrarily large by replacing the system $\{S_i, p_i\}_{i \in \mathcal{A}}$ by its iteration, we obtain the main theorem.

References

- [1] Pertti Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995. Fractals and rectifiability. MR1333890 (96h:28006)
- [2] Robert S. Strichartz, *Self-similar measures and their Fourier transforms. I*, Indiana Univ. Math. J. **39** (1990), no. 3, 797–817. MR1078738 (92k:42015)
- [3] Lai-Sang Young, *Large deviations in dynamical systems*, Trans. Amer. Math. Soc. **318** (1990), no. 2, 525–543. MR975689 (90g:58069)
- [4] Y. Takahashi, *Asymptotic behaviours of measures of small tubes: entropy, Liapunov's exponent and large deviation*, Dynamical systems and applications (Kyoto, 1987), World Sci. Adv. Ser. Dynam. Systems, vol. 5, World Sci. Publishing, Singapore, 1987, pp. 1–21. MR974155 (90a:58098)
- [5] S. R. S. Varadhan, *Large deviations*, Ann. Probab. **36** (2008), no. 2, 397–419, DOI 10.1214/07-AOP348. MR2393987 (2009d:60070)